

NL
PH
030933
PCTMAT.
DOSSIERWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

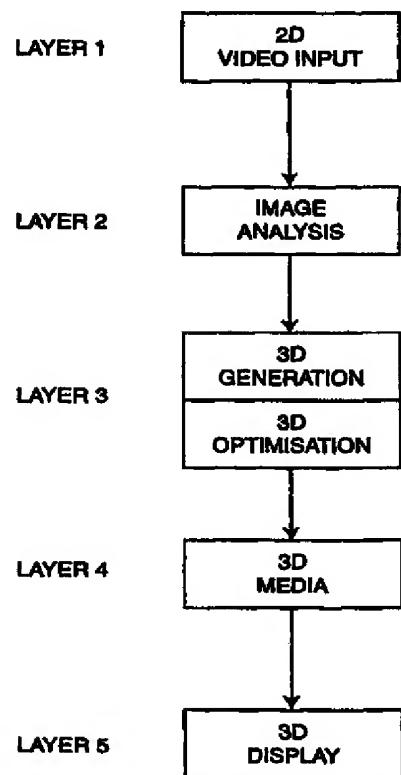
(51) International Patent Classification ⁶ :		(11) International Publication Number:	WO 99/12127
G06T 15/00, H04N 13/00, 15/00	A1	(43) International Publication Date:	11 March 1999 (11.03.99)

(21) International Application Number:	PCT/AU98/00716	(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date:	2 September 1998 (02.09.98)	
(30) Priority Data:		
PO 8944	2 September 1997 (02.09.97)	AU
(71) Applicant (for all designated States except US): XENOTECH RESEARCH PTY. LTD. [AU/AU]; Suite 1 & 2, 41 Walters Drive, Osborne Park, W.A. 6017 (AU).		
(72) Inventor; and		
(75) Inventor/Applicant (for US only): HARMAN, Philip, Victor [AU/AU]; 45 Ventor Street, Scarborough, W.A. 6019 (AU).		
(74) Agent: WATERMARK PATENT & TRADEMARK ATTORNEYS; "Durack Centre", 4th floor, 263 Adelaide Terrace, Perth, W.A. 6000 (AU).		

(54) Title: IMAGE PROCESSING METHOD AND APPARATUS

(57) Abstract

An image conversion system for converting monoscopic images for viewing in three dimensions including: an input means (1) adapted to receive the monoscopic images; a preliminary analysis means to determine if there is any continuity between a first image and a second image of the monoscopic image sequence; a secondary analysis means (2) for receiving monoscopic images which have a continuity, and analysing the images to determine the speed and direction of motion, and the depth, size and position of objects; a first processing means (3) for processing the monoscopic images based on data received from the preliminary analysis means or the secondary analysis means; a second processing means capable of further processing images received from the first processing means; a transmission means (4) capable of transferring the processed images to a stereoscopic display system (5).



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

IMAGE PROCESSING METHOD AND APPARATUS

FIELD OF INVENTION

The present invention relates generally to stereoscopic image systems, and in particular to the synthesis of stereoscopic image pairs from monoscopic images for stereoscopic display. The present invention may also be directed towards a five layer method for producing stereoscopic images, that digitises a monoscopic source, analyses it for motion, generates the stereoscopic image pairs, optimises the stereoscopic effect, transmits or stores them and then enables them to be displayed on a stereoscopic display device.

BACKGROUND ART

The advent of stereoscopic or three dimensional (3D) display systems which create a more realistic image for the viewer than conventional monoscopic or two dimensional (2D) display systems, requires that stereoscopic images be available to be seen on the 3D display systems. In this regard there exists many monoscopic image sources, for example existing 2D films or videos, which could be manipulated to produce stereoscopic images for viewing on a stereoscopic display device.

Preexisting methods to convert such monoscopic images for stereoscopic viewing do not produce acceptable results. Other attempts in film and video have used techniques to duplicate the stereoscopic depth cue of "Motion Parallax". These involve producing a delay for the images presented to the trailing eye when lateral, left or right, motion is present in the images. Other attempts have used 'Lateral Shifting' of the images to the left and right eyes to provide depth perception.

However, these two techniques are limited and generally only suit specific applications. For example, the Motion Parallax technique is only good for scenes with left or right motion and is of limited value for the stereoscopic enhancement of still scenes. The Lateral Shifting technique will only give an overall depth effect to a scene and will not allow different objects at varying depths to be perceived at the depths where they occur. Even the combination of these two techniques will only give a limited stereoscopic effect for most 2D films or videos.

Some existing approaches demonstrate limitations of these techniques. When an image has vertical motion and some lateral motion and a delay is provided to the image presented to the trailing eye then the result is often a large vertical disparity between the left and right views such that the images are
5 uncomfortable to view. Scenes with contra motion, such as objects moving left and right in the same scene are also uncomfortable to view. Certain embodiments of these methods define that when objects of varying depths are present in an image there is a distinct 'card board cut-out' appearance of the objects with distinct depth layers rather than a smooth transition of objects from
10 foreground to background.

In all these approaches no successful attempt has been made to develop a system or method to suit all image sequences or to resolve the problem of viewer discomfort or to optimise the stereoscopic effect for each viewer or display device.

15 **OBJECTS OF THE INVENTION**

There is therefore a need for a system with improved methods of converting monoscopic images into stereoscopic image pairs and a system for providing improved stereoscopic images to a viewer.

An object of the present invention is to provide such a system with
20 improved methods.

SUMMARY OF INVENTION

In order to address the problems noted above the present invention provides in one aspect a method of manipulating monoscopic images into stereoscopic image pairs, including the steps of: analysing the monoscopic
25 images to determine the nature of motion within an image; comparing any detected motion with a predefined range of motion categories; processing the monoscopic images using at least one processing method depending on the motion category to form stereoscopic image pairs.

In another aspect, the present invention provides a stereoscopic system
30 including: an input means capable of receiving monoscopic images; an analysis means for analysing the images to determine an image category; a conversion means capable of converting the monoscopic images into stereoscopic image

pairs as a function of the selected image category for stereoscopic viewing.

Ideally, the input means also includes a means to capture and digitise the monoscopic images.

Preferably the image analysis means is capable of determining the speed
5 and direction of motion, the depth, size and position of objects and background
within an image.

In a further aspect the present invention provides a method of optimising
the stereoscopic image to further improve the stereoscopic effect and this
process is generally applied prior to transmission, storage and display.

10 In yet a further aspect the present invention provides a method of
improving stereoscopic image pairs by adding a viewer reference point to the
image.

15 In still yet a further aspect the present invention provides a method of
analysing monoscopic images for conversion to stereoscopic image pairs
including the steps of: scaling each image into a plurality of regions; comparing
each region of a first image with corresponding and adjacent regions of a
second image to determine the nature of movement between said first image
and said second image.

20 Preferably a motion vector is defined for each image based on a
comparison of the nature of motion detected with predefined motion categories
ranging from no motion to a complete scene change.

25 In yet a further aspect the present invention provides an image
conversion system for converting monoscopic images for viewing in three
dimensions including: an input means adapted to receive monoscopic images;
a preliminary analysis means to determine if there is any continuity between a
first image and a second image of the monoscopic image sequence; a secondary analysis means for receiving monoscopic images which have a
continuity, and analysing the images to determine at least one of the speed and
direction of motion, of the depth, size and position of objects; a first processing
30 means for processing the monoscopic images based on data received from the
preliminary analysis means of the secondary analysis means.

Preferably a second processing means capable of further processing

images received from the first processing means is included. Ideally the system would also include a transmission means capable of transferring the processed images to a stereoscopic display system or storage means.

Preferably a third processing means is provided for optionally enhancing 5 the images received from the second processing means prior to transmitting the converted images to the stereoscopic display device.

In yet a further aspect the present invention provides a method for converting monoscopic images for viewing in three dimensions including: a first layer adapted to receive a monoscopic image; a second layer adapted to 10 receive the monoscopic image and analyse the monoscopic image to create image data; a third layer adapted to create stereoscopic image pairs from the monoscopic image using at least one predetermined technique selected as a function of the image data; a fourth layer adapted to transfer the stereoscopic image pairs to a stereoscopic display means; a fifth layer consisting of a 15 stereoscopic display means.

Preferably the first layer is further adapted to convert any analogue images into a digital image. Also, the second layer is preferably adapted to detect any objects in a scene and make a determination as to the speed and direction of any such motion. Conveniently, the image may be compressed prior 20 to any such analysis.

Preferably the third layer further includes an optimisation stage to further enhance the stereoscopic image pairs prior to transmitting the stereoscopic image pairs to the stereoscopic display means. Conveniently, the fourth layer may also include a storage means for storing the stereoscopic image pairs for 25 display on the stereoscopic display means at a later time.

ADVANTAGES

It will be appreciated that the process of the present invention can be suspended at any stage and stored for continuation at a later time or transmitted for continuation at another location if required.

30 The present invention provides a conversion technology with a number of unique advantages including:

1) Realtime or Non-realtime conversion

The ability to convert monoscopic images to stereoscopic image pairs can be performed in realtime or non-realtime. Operator intervention may be applied to manually manipulate the images. An example of this is in the conversion of films or videos where every sequence may be tested and 5 optimised for its stereoscopic effect by an operator.

2) Techniques include stereoscopic enhancement

The present invention utilises a plurality of techniques to further enhance the basic techniques of motion parallax and lateral shifting (forced parallax) to generate stereoscopic image pairs. These techniques include but are not 10 limited to the use of object analysis, tagging, tracking and morphing, parallax zones, reference points, movement synthesis and parallax modulation techniques.

3) Detection and correction of Reverse 3D

Reverse 3D is ideally detected as part of the 3D Generation process by 15 analysing the motion characteristics of an image. Correction techniques may then employed to minimise Reverse 3D so as to minimise viewer discomfort.

4) Usage in all applications - includes transmission and storage

The present invention discloses a technique applicable to a broad range 20 of applications and describes a complete process for applying the stereoscopic conversion process to monoscopic applications. The present invention describes on the one hand techniques for 3D Generation where both the image processing equipment and stereoscopic display equipment are located substantially at the same location. While on the other hand techniques are 25 defined for generation of the stereoscopic image pairs at one location and their transmission, storage and subsequent display at a remote location.

5) Can be used with any stereoscopic display device

The present invention accommodates any stereoscopic display device and ideally has built in adjustment facilities. The 3D Generation process can 30 also take into account the type of display device in order to optimise the stereoscopic effect.

BRIEF DESCRIPTION OF FIGURES

The invention will be more fully understood from the following description of a preferred embodiment of the conversion method and integrated system and as illustrated in the accompanying figures. It is, however, to be appreciated that the present invention is not limited to the described embodiment.

5 Figure 1 shows the breakdown into layers of a complete system utilising the present invention.

Figure 2 shows a possible use of multiple processors with a complete system utilising the present invention.

10 Figure 3 shows a flow diagram of Layer 1 (Video Digitising) and the first part of Layer 2 (Image Analysis).

Figure 4 shows the second part of a flow diagram of Layer 2.

Figure 5 shows the third part of a flow diagram of Layer 2.

Figure 6 shows the fourth part of a flow diagram of Layer 2.

Figure 7 shows a flow diagram of the first part of Layer 3 (3D Generation).

15 Figure 8 shows the second part of a flow diagram of Layer 3 and Layer 4 (3D Media - Transmission & Storage) and Layer 5 (3D Display).

DETAILED DESCRIPTION

The present invention aims to provide a viewer with a stereoscopic image that uses the full visual perception capabilities of an individual. Therefore it is 20 necessary to provide the depth cues the brain requires to interpret such images.

INTRODUCTION

Humans see by a complex combination of physiological and psychological processes involving the eyes and the brain. Visual perception involves the use of short and long term memory to be able to interpret visual 25 information with known and experienced reality as defined by our senses. For instance, according to the Cartesian laws on space and perspective the further an object moves away from the viewer the smaller it gets. In other words, the brain expects that if an object is large it is close to the viewer and if it is small it is some distance off. This is a learned process based on knowing the size of the 30 object in the first place. Other monoscopic or minor depth cues that can be represented in visual information are for example shadows, defocussing, texture, light, atmosphere,.

These depth cues are used to great advantage in the production of 'Perspective 3D' video games and computer graphics. However, the problem with these techniques in achieving a stereoscopic effect is that the perceived depth cannot be quantified: it is an illusion of displaying 2D objects in a 2D environment. Such displays do not look real as they do not show a stereoscopic image because the views to both eyes are identical.

DEPTH CUES

Stereoscopic images are an attempt to recreate real world visuals, and require much more visual information than 'Perspective 3D' images so that 10 depth can be quantified. The stereoscopic or major depth cues provide this additional data so that a person's visual perception can be stimulated in three dimensions. These major depth cues are described as follows :-

1) **Retinal Disparity** - refers to the fact that both eyes see a slightly different view. This can easily be demonstrated by holding an object in front of a 15 person's face and focussing on the background. Once the eyes have focused on the background it will appear as though there are actually two objects in front of the face. Disparity is the horizontal distance between the corresponding left and right image points of superimposed retinal images. While Parallax is the actual spatial displacement between the viewed images.

20 2) **Motion Parallax** - Those objects that are closer to the viewer will appear to move faster even if they are travelling at the same speed as more distant objects. Therefore relative motion is a minor depth cue. But the major stereoscopic depth cue of lateral motion is the creation of motion parallax. With motion in an image moving from right to left, the right eye is the leading eye 25 while the left eye becomes the trailing eye with its image being delayed. This delay is a normal function of our visual perception mechanism. For left to right motion the right eye becomes the trailing eye. The effect of this delay is to create retinal disparity (two different views to the eyes), which is perceived as binocular parallax thus providing the stereoscopic cue known as Motion 30 Parallax.

3) **Accommodation** - The eye brings an object into sharp focus by either compressing the eye lens (more convex shape for close object) or

expanding the eye lens (less convex shape for far object) through neuromotor activity. The amount and type of neuromotor activity is a stereoscopic cue for depth in an image.

4) **Convergence** - Is the response of the eye's neuromotor system that
5 brings images of an object into alignment with the central visual area of the eyes such that only one object is seen. For example, when a finger held at arms length is viewed by both eyes and slowly brought towards the face, the eyes turn inwards (converge) indicating that the finger is getting closer. That is, the depth to the finger is decreasing.
10 The eyes convergence response is physiologically linked to the accommodation mechanism in normal vision. In stereoscopic viewing, when viewers are not accommodated to the 'Fixation Plane' (that to which the eyes are converged), they may experience discomfort. The 'Plane of Fixation' is normally the screen plane.

15 **OVERVIEW - 5 LAYER APPROACH**

The present invention describes a system that is capable of taking any monoscopic input and converting it to an improved stereoscopic output. For ease of description this complete system can be broken down into a number of independent layers or processes, namely:-

20 LAYER 1 - Monoscopic Image Input (typically video input)
LAYER 2 - Image Analysis
LAYER 3 - 3D Generation
LAYER 4 - 3D Media (Transmission or Storage)
LAYER 5 - 3D Display

25 Figure 1 shows this top down approach to the stereoscopic conversion process, where video or some other monoscopic image source is input, images are analysed, stereoscopic image pairs are generated, transmitted and/or stored and then displayed on a stereoscopic display. Each Layer describes an independent process of the complete system from monoscopic image input to
30 stereoscopic display. However, it will be appreciated that the various layers may be operated independently.

APPLICATIONS

Generally, all five layers are used, from monoscopic image input to display for a particular application. For example, this system may be used in theatres or cinemas. In such an application the 2D video input can take the form of analogue or digital to the video sources. These sources would then be
5 analysed to determine speed and direction of any motion. The processes would then work in either real-time or non real-time in order to create the 3D images. This can be further optimised through the use of borders, parallax modification, reverse 3D analysis, shading, and/or texturing. The 3D images may then be stored or transmitted to a 3D display, comprising shutterglasses, polarising
10 glasses or an autostereoscopic display.

This system may also be adapted for use with cable or pay-TV systems. In this application the 2D video input could be video from a VTR, a laser disc, or some other digital source. Again the 3D Generation and/or optimisation can proceed in either real time or non real time. The 3D media layer would
15 conveniently take the form of transmission via cable or satellite to enable 3D display on TV, video projector, or an auto stereoscopic display.

The system may also be used with video arcade games, in multimedia, or with terrestrial or network TV. Depending on the application the 2D video input layer may obtain source monoscopic images from a games processor, video
20 from a laser disc, video from VTR, video from a network, or some other digital storage device or digital source or telecine process. The 3D Generation can take place in real time or non real time, and be generated by computer at a central conversion site, in a user's computer, on a central processor, or some other image processor. The stereoscopic images can then be stored on video
25 or other digital storage device, prior to distribution to cinemas or transmission by a local network. These stereoscopic images may also be transmitted to video projectors via a local transmission, or alternatively via VHF/UHF facilities or satellite.

The 3D display is dependent on the application required, and can take
30 the form of an auto stereoscopic display device, a video projector with polarising glasses, a local monitor with shutter-glasses, or a set-top box with suitable viewing glasses.

Single & Multiple Processors

The complete system can be operated on a single processor with all five layers being processed together or individually in realtime or non-realtime (Layers 2, 3 and 4). Layers 2 and 3 can be further segmented to suit a
5 multitasking or multiprocessor environment, as can be seen in Figure 2 for example.

The use of multiple processors can also be configured to the application on hand. For example, layers 1 and 2 could be handled by a first processor, and layers 3 to 5 by a second processor. If desired, the first processor of this
10 arrangement could be used as a look-ahead processor, and the second processor could generate the stereoscopic images after a delay. Alternatively, a first processor could be used to receive realtime video, digitise the video and transfer the digitised video to a suitable digital storage device. A second processor, either on site or remotely, could then analyse the digitised image and
15 perform the necessary tasks to display a stereoscopic image on a suitable display device.

Look-ahead processing techniques may be employed to predict trends in sequences of film or video so that the image processing modes may be more efficiently selected to optimise the overall stereoscopic effect.

20 The present invention is primarily concerned with the analysis of monoscopic images and conversion of the monoscopic images into stereoscopic image pairs together with the optimisation of the stereoscopic effect. In this regard the present invention is applicable to a broad range of monoscopic inputs, transmission means and viewing means. However, for
25 completeness all five defined layers will be described herein:

LAYER 1 - IMAGE OR VIDEO INPUT

Layer 1 requires that a monoscopic image source or video input is provided. This source may be provided as either a digital image source or an analogue image source which may then be digitised. These image sources
30 may include:-

1) Analogue Source

- a) Tape based - VCR / VTR or Film.
- b) Disk based - Laser Disk.
- 5 c) Video Camera or other realtime image capture device.
- d) Computer generated images or graphics.

2) Digital Source

- a) Tape based - Typical examples are DAT, AMPEX's DCT, SONY's Digital Betacam, Panasonic's digital video formats or the new Digital 10 Video Cassette (DVC) format using 6.5mm tape.
- b) Disk based storage - Magneto Optical (MO) hard disk (HD), compact disk (CD), Laser Disk, CD-ROM, DAT, Digital Video Cassette (DVC) or Digital Video Disk (DVD) based data storage devices - uses JPEG, MPEG or other digital formats
- 15 c) Video Camera or other realtime image capture device
- d) Computer generated images or graphics.

What is important for the conversion process of the present invention is that a monoscopic image source be provided. It is noted that a stereoscopic image source may be provided which would generally obviate the need for 20 layers 1 to 3, however, any such stereoscopic image may be passed through an optimisation stage prior to display.

LAYER 2 - IMAGE ANALYSIS

Referring now to Figures 3 to 8 which show flow diagrams demonstrating a preferred arrangement of the present invention.

25 Following reception of 2D images, digitised video or digital image data is processed on a field by field or image by image basis in realtime or non-realtime by hardware, software or by a combination of both. Firstly, the image analysis process occurs including the steps of : -

- 1) Image compression
- 30 2) Motion detection
- 3) Object detection
- 4) Motion analysis

1) Image Compression

Compression of the image is not essential, however, for many processes and applications, compression is a practical option particularly, where the processor is not powerful enough to process a full resolution image in the time required.

Preferably the images are scaled to smaller dimensions. The scaling factor is dependent on the digital video resolution used for each image, and is usually defined by the type of image capture facility used in the digitising process.

10 2) Motion Detection

In a preferred embodiment each image may be analysed in blocks of pixels. A motion vector is calculated for each block by comparing blocks from one image with corresponding blocks from an adjacent image that are offset horizontally and/or vertically by up to a predetermined number of pixels, for example ±9, and recording the position that gives the minimum Mean Squared Error.

For each block, the vector and minimum and maximum Mean Squared Error are recorded for later processing.

To save on processing time, vectors need not be calculated if there is no detail in the block, for example, when the block is a homogeneous colour.

Other methods for calculating the motion can be utilised, for example image subtraction. The present embodiment uses the Mean Squared Error method.

3) Object Detection

25 An Object is defined as a group of pixels or image elements that identify a part of an image that has common features. Those characteristics may relate to regions of similar luminance value (similar brightness), chrominance value (similar colour), motion vector (similar speed and direction of motion) or similar picture detail (similar pattern or edge).

30 For example a car driving past a house. The car is a region of pixels or pixel blocks that is moving at a different rate to the background. If the car stopped in front of the house then the car would be difficult to detect, and other

methods may be used.

A connectivity algorithm may be used to combine the motion vectors into regions of similar motion vectors. An Object may be comprised of one or more of such regions. Other image processing algorithms, such as edge detection
5 etc, may be used in the detection of Objects.

Once Objects are identified in an image they are preferably tagged or given an identification number. These Objects and their relevant details (for example position, size, motion vector, type, depth) are then stored in a database so that further processing may occur. If an Object is followed over a
10 sequence of images then this is known as Object Tracking. By tracking Objects and analysing their characteristics they can be identified as being foreground or background Objects and therefore enhanced to emphasise their depth position in an image.

4) Motion Analysis

15 Once Objects have been detected, the Objects can be analysed to determine the overall speed and direction of motion in the image. In the preferred embodiment, this stage determines the type of motion in the image, and also provides an overall vector.

By using the Object Detection information and comparing the data to
20 several image motion models a primary determination can be made as to the best method to convert monoscopic images to stereoscopic image pairs.

The image motion models as used in the preferred embodiment of the present invention are: -

- a) Scene Change
- 25 b) Simple Pan
- c) Complex Pan
- d) Moving Object over stationary background
- e) Foreground Object over moving background
- f) No Motion

30 Other motion models may be used as required.

a) Scene Change

A scene change as the name suggests is when one image has little or no

commonality to a previous image or scene. It may be detected as a very large absolute difference in luminance between the two images, or a large difference in the colours of the two images.

In a preferred arrangement a scene change may be determined when the median of the differences of luminance values (0-255) between previous and current images is typically above 30. This value may vary with application but trial and error has determined that this value is appropriate for determining most scene changes.

A secondary test to determine a scene change can be when there are too many regions of motion vectors, which appears like random noise on the image and is likely due to a scene change. This may occur if there is a very large amount of motion in the image.

A third technique to detect a scene change is to analyse the top few lines of each image to detect a scene change. The top of each image changes the least.

Alternatively, when the majority of motion vector blocks have large error values the difference between the two images is too great and will therefore be considered as a scene change.

Scene change and Field Delay

In the preferred embodiment when there is lateral motion detected in a scene the image to the trailing eye is delayed by an amount of time that is inversely proportional to the speed of the motion. For an image moving right to left the trailing eye is the left eye and for an image moving left to right the trailing eye is the right eye.

The image sequence delay (or Field Delay) to the trailing eye, may be created by temporally delaying the sequence of video fields to the trailing eye by storing them in digital form in memory. The current video field is shown to the leading eye and the delayed image to the trailing eye is selected from the stored video fields depending on the speed of the lateral motion.

Over a number of fields displayed, a history as to the change in motion and change in Field Delays to the trailing eye can be maintained. This helps in smoothing the stereoscopic effect by enabling the image processor to predict

any motion trends and to react accordingly by modifying the delay so that there are no sudden changes.

If a scene change is detected the Field Delay for the preferred embodiment of the present invention is set to zero to prevent the image breaking apart and the Field Delay history is also reset. Field Delay history is preferably reset on each scene change.

b) Simple Pan

A simple pan describes a lateral motion trend over a series of images whereby the majority of analysed motion is in one direction. This will preferably also cover the situation where the majority of the scene has a consistent motion, and no stationary objects are detected in the foreground.

A simple pan can be detected as the major Object having a non zero motion vector..

The result of a simple pan is that a positive motion vector is generated if the scene is moving to the right (or panning left). In this case, the image to the right eye will be delayed. Similarly, a negative motion vector is generated if the scene is moving to the left (or panning right). In this case, the image to the left eye will be delayed.

c) Complex Pan

A complex pan differs from a simple pan in that there is significant vertical motion in the image. Therefore, in the preferred embodiment, to minimise vertical disparity between the stereoscopic image pair sequences, Field Delay is not applied and only Object Processing is used to create a stereoscopic effect. Field Delay history is stored to maintain continuity with new lateral motion.

d) Moving Object over Stationary Background

A moving object over a stationary background is simply the situation whereby the majority of a scene has no motion, and one or more moving Objects of medium size are in scene. This situation also results in a positive motion vector if the majority of Objects are moving to the right, and a negative motion vector if the majority of Objects are moving to the left. A positive motion vector produces a delay to the right eye, and a negative motion vector produces a delay to the left eye. .

In the case where the motion vectors of the Objects in the scene are not consistent, for example, objects moving to the left and right in the same scene, then Contra Motion exists and Reverse 3D correction techniques may be applied.

5 **e) Foreground Object over Moving Background**

A Foreground Object over a moving background refers to the situation where a majority of the scene has motion, and an Object having a different motion is in the scene, for example a camera following a person walking. A Background Object is detected as a major Object of non-zero motion vector

10 (That is, a panning background) behind an Object of medium size with zero or opposite motion vector to the main Object, or a major Object of zero vector in front of minor Objects of non zero vector that are spread over the entire field (That is, a large stationary object filling most of the field, but a pan is still visible behind it).

15 A decision should be made as to whether the foreground Object should be given priority in the generation of Motion Parallax, or whether the background should be given priority. If the background contains a large variation in depth (for example, trees), then motion vectors are assigned as if a Simple pan was occurring. If the background contains little variation in depth
20 (for example, a wall) then a motion vector is assigned that is antiparallel or negative. .

25 When the background contains a large variation in depth, and a motion vector is assigned to the scene as per Simple Pan methods, then the foreground object will be in Reverse 3D, and suitable correction methods should be applied.

f) No Motion

If no motion is detected such that the motion vectors are entirely zero, or alternatively the largest moving Object is considered too small, then the Field Delay will be set to zero. This situation can occur where only random or noise
30 motion vectors are determined, or where no motion information is available, for example during a pan across a blue sky.

LAYER 3 - 3D GENERATION

Once the images are analysed they can then be processed to create the stereoscopic image pairs.

When viewing a real world scene both eyes see a slightly different image. This is called retinal disparity. This in turn produces stereopsis or depth perception. In other words we see stereoscopically by having each eye see a slightly different image of the same scene.

Parallax on the other hand is defined as the amount of horizontal or lateral shift between the images which is perceived by the viewer as retinal disparity. When a stereoscopic image pair is created, a three-dimensional scene is observed from two horizontally-shifted viewpoints.

The present invention utilises a number of image and object processing techniques to generate stereoscopic image pairs from monoscopic images. These techniques include:

- 1) Motion Parallax
- 2) Forced Parallax (Lateral Shifting)
- 3) Parallax Zones
- 4) Image Rotation about the Y-Axis
- 5) Object Processing

20 1) Motion Parallax

When a scene is moving from right to left, the right eye will observe the scene first while the left eye will receive a delayed image and visa versa for a scene moving in the opposite direction. The faster the motion the less delay between the images to both eyes. This is known as motion parallax and is a major depth cue. Therefore, if there is lateral motion in a scene, by creating a delay between the images to the eyes a stereoscopic effect will be perceived.

a) Field Delay Calculation

Once the nature of the motion in an image has been analysed and an overall motion vector determined, the required Field Delay can then be calculated. Preferably, the calculated Field Delay is averaged with previous delays to filter out 'noisy' values and also to prevent the Field Delay changing too quickly.

As stated above, the faster the motion the less delay between the image to each eye. Accordingly, smaller values of Field Delay are used in scenes with large motion vectors, whereas larger delays are used in scenes with little lateral motion. That is, an inverse relationship exists in the preferred embodiment 5 between the delay and amount of motion.

When a scene change is determined, the history of Field Delays should be reset to zero, as if no motion had occurred previously. At the first detection of motion when a non zero Field Delay is calculated whilst the history of Field Delays is still zero, the entire history of Field Delay is set to the calculated Field 10 Delay. This enables the system to immediately display the correct Field Delay when motion is detected.

b) Field Delay Implementation

Motion Parallax can be generated in hardware and software by storing digitised images in memory. Preferably, the digitised images could be stored in 15 a buffer and a single input pointer used with two output pointers, one for the left eye image and one for the right eye image. The leading eye's image memory pointer is maintained at or near the current input image memory pointer while the delayed eyes image memory pointer is set further down the buffer to produce a delayed output. Many images may be stored, up to 8-10 video fields 20 is typical in video applications. The delay is dependent on the speed of the motion analysed in the image. Maximum field delay is when there is minimum motion.

2) Forced Parallax (Lateral Shifting)

Forced parallax can be created by introducing a lateral shift between :-

- 25 i) An exact copy of an image and itself
- ii) The two fields of a video frame
- iii) Two frames of a film sequence
- iv) A transformed copy of an image and its original

A Negative lateral shift is produced by displacing the left image to the 30 right and the right image to the left by the same amount (establishes a depth of field commencing from the screen plane and proceeding in front of it) and a Positive lateral shift by displacing the left image to the left and the right image to

the right by the same amount (establishes a depth of field commencing from the screen plane and receding behind it).

Forced Parallax may be reduced to enhance the stereoscopic effect for a stationary object in front of a pan, where the object is 'placed' closer to the 5 screen plane and the background is 'pushed back' from the defined object plane.

3) Parallax Zones

Because most scenes are viewed with the background at the top and the foreground at the bottom it is possible to accentuate a scene's depth by 'Veeing' 10 the Forced Parallax. This is done by laterally shifting the top of the image more than the bottom of an image thus accentuating the front to back depth observed in a scene.

Another technique is to use a combination of Motion Parallax and Forced 15 Parallax on different parts of the image. For example, by splitting the image vertically in half and applying different parallax shifts to each side, a scene such as looking forwards from a moving train down a railway track has the correct stereoscopic effect. Otherwise one side would always appear in Reverse 3D.

4) Image Rotation about the Y-Axis

When an object is moving towards the viewer in a real world scene, the 20 object is rotated slightly in the view for each eye. The rotation effect is more pronounced as the object moves closer. Translating this rotation into the stereoscopic image pairs defines the effect as follows :-

i) Moving towards the viewer - The left image is rotated vertically about its central axis in an anti-clockwise direction and the right image in a clockwise 25 direction.

ii) Moving away from the viewer - The left image is rotated vertically about its central axis in a clockwise direction and the right image in an anti-clockwise direction.

Therefore, by image rotation, the perspective of objects in the image is 30 changed slightly so that depth is perceived. When this technique is combined with Forced Parallax for certain scenes the combined effect provides very powerful stereoscopic depth cues.

5) Object Processing

Object processing is performed to further enhance the stereoscopic effect, particularly in still images, by separating the Objects and background so that these items can be processed independently. It is most effective when the
5 objects are large in size, few in number and occupy distinct depth levels throughout the depth of field.

A database for Object Tagging and Object Tracking can be used to establish trends so that an Object can be digitally 'cut out' from its background and appropriate measures taken to enhance the stereoscopic effect. Once
10 processing has taken place the Object is 'Pasted' back in the same position on to the background again. This can be termed the 'Cut and Paste' technique and is useful in the conversion process.

By integrating the processes of Object Tagging, Tracking, Cutting and Pasting a powerful tool is available for enabling Object Processing and
15 Background Processing.

Another Object Processing technique is Object Layering which defines an independent depth layer for each moving Object. The Object can then be placed anywhere on an image because the background fill detail has been defined when the Object was not in that position. This is not generally possible
20 with a still Object unless the background fill-in is interpolated.

A most important issue in stereoscopic conversion is the correction of Reverse 3D and Accommodation/Convergence imbalances that cause viewer discomfort. Object Processing in the preferred embodiment allows corrections to this problem too.

25 **a) Mesh Distortion and Morphing** - This Object processing technique allows an Object to be cut and pasted onto a distorted mesh to enhance depth perception. By distorting an Object in the left eye image to the right and by distorting the same object in the right eye image to the left, thus creating Object Parallax, the Object can be made to appear much closer to a
30 viewer when using a stereoscopic display device.

b) Object Barrelling - This technique is a specific form of Mesh Distortion and refers to a technique of cutting an Object from the image and

wrapping onto a vertically positioned half barrel. This makes the Object appear to have depth by making the centre portion of the Object appear closer than the Object edges.

c) **Object Edge Enhancement** - By emphasising the edges of an Object there is greater differentiation between the background or other Objects in an image. The stereoscopic effect is enhanced in many applications by this technique.

d) **Object Brightness Enhancement** - In any image the eye is always drawn to the largest and brightest objects. By modifying an Object's luminance the Object can be emphasised more over the background, enhancing the stereoscopic effect.

e) **Object rotation about Y-axis** - Object rotation about the Y-axis refers to a similar process to that of image rotation about the Y-axis, except that this time the rotation occurs to the Object only. If the Object in the stereoscopic image pair is 'Cut' from its background and rotated slightly the change in perspective generated by the rotation is perceived as depth.

3D OPTIMISATION

1) Reference Points or Borders

When using a normal TV or video monitor to display stereoscopic images the eye continually observes the edge of the monitor or screen and this is perceived as a point of reference or fixation point for all depth perception. That is, all objects are perceived at a depth behind or in front of this reference point.

If the edge of the monitor is not easily seen because of poor ambient lighting or due to its dark colour then this reference point may be lost and the eyes may continually search for a fixation point in the 3D domain. Under prolonged stereoscopic viewing this can cause eye fatigue and decreased depth perception. A front or rear projection screen display system may also suffer from the same problems.

The present invention therefore preferably also defines a common border or reference point within a viewed image. Ideally the reference plane is set at the screen level and all depth is perceived behind this level. This has the advantage of enhancing the stereoscopic effect in many scenes.

This reference point can be a simple video border or reference graphic and, for example, may be of the following types :

- i) A simple coloured video border around the perimeter of the image.
- ii) A complex coloured video border consisting of two or more concentric borders that may have opaque or transparent sections between them. For example, a 2-3cm wide mesh border or a wide outer border with two thin inner borders.
- iii) A partial border that may occupy any one edge, or any two horizontal or vertical edges.
- 10 iv) A LOGO or other graphic located at some point within the image.
- v) A picture within a picture.
- vi) A combination of any of the above.

What is essential in this embodiment is that the eyes of the viewer be provided with a reference point by which the depth of the objects in the image 15 can be perceived.

If a border or graphic is added at the 3D Generation level then it may be specified to provide a reference point at a particular depth by creating left and right borders that are laterally shifted from each other. This enables the reference or fixation point to be shifted in space to a point somewhere behind or 20 in front of the screen level. Borders or graphics defined with no parallax for the left and right eyes will be perceived at the screen level. This is the preferred mode of the present invention.

A image border or reference graphic may be inserted at the 3D Generation point or it may be defined externally and genlocked onto the 25 stereoscopic image output for display. Such an image border or reference graphic may be black, white or coloured, plain or patterned, opaque, translucent or transparent to the image background, or it may be static or dynamic. Whilst a static border is appropriate in most instances, in some circumstances a moving or dynamic border may be used for motion enhancement.

30 **2) Parallax Adjustment - Depth Sensitivity Control**

Stereoscopic images viewed through a stereoscopic display device automatically define a depth range (called depth acuity) which can be increased

or decreased by modifying the type and amount of parallax applied to the image or objects. It has been found that different viewers have varying stereoscopic viewing comfort levels based on the depth range or amount of stereopsis defined by stereoscopic image pairs. That is, while some viewers prefer a 5 pronounced stereoscopic effect with a greater depth range, others prefer an image with minimal depth.

To adjust the level of depth sensitivity and viewing comfort many techniques may be used, namely :

- i) Varying the amount of Motion Parallax by varying the Field Delay
- 10 ii) Varying the amount of Forced Parallax to an image
- iii) Varying the amount of Parallax applied to objects

By reducing the maximum level of Parallax the depth range can be reduced, improving the viewing comfort for those with perception faculties having greater sensitivity to stereoscopy.

15 3) Parallax Smoothing

Parallax Smoothing is the process of maintaining the total amount of Parallax (Motion Parallax plus Forced Parallax) as a continuous function. Changes in Field Delay for specific motion types, that is, Simple Pan and Foreground Object Motion, cause discontinuities in the amount of Motion 20 Parallax produced, which are seen as "jumps" in the stereoscopic images by the viewer. Discontinuities only occur in the image produced for the trailing eye, as the leading eye is presented with an undelayed image . These discontinuities can be compensated for by adjusting the Forced Parallax or Object Parallax in an equal and opposite direction for the trailing eye, thus 25 maintaining a continuous total parallax.

The Forced Parallax or Object Parallax is then adjusted smoothly back to its normal value, ready for the next change in Field Delay. The adjustments made to Forced Parallax by Parallax Smoothing are a function of Field Delay change, motion type and motion vector. To implement Parallax Smoothing, the 30 Forced Parallax for the left and right eye images should be independently set.

4) Parallax Modulation

The Forced Parallax technique of creating a stereoscopic effect can also be used to moderate the amount of stereopsis detected by the viewer. This is 5 done by varying the Forced Parallax setting between a minimum and maximum limit over a short time such that the perceived depth of an object or image varies over time. Ideally the Forced Parallax is modulated between its minimum and maximum settings every 0.5 to 1 second. This enables a viewer to accommodate to their level of stereoscopic sensitivity.

10 5) Movement Synthesis

By creating pseudo movement, by randomly moving the background in small undetectable increments, the perceived depth of foreground objects is emphasised. Foreground objects are 'Cut' from the background, the background is altered pseudo-randomly by one of the techniques below and 15 then the foreground object is 'Pasted' back on to the background ready for display. Any of the following techniques may be used :-

- 20 i) Luminance values varied on a pseudo-random basis
- ii) Chrominance values varied on a pseudo-random basis
- iii) Adding pseudo-random noise to the background to create movement

6) Reverse 3D analysis and correction

Reverse 3D occurs when the depth order of Objects created by Parallax is perceived to be different to that corresponding to the depth order in the real world. This generally leads to viewer discomfort and should be corrected. 25 When converting monoscopic images to stereoscopic image pairs Reverse 3D may be produced by :-

- i) Contra motion, objects moving left and right in the same image.
- ii) Objects and background moving in different directions.
- iii) Many objects moving at varying speeds

30 Reverse 3D is corrected by analysing the nature of the motion of the objects in an image and then manipulating each Object individually using mesh distortion techniques so that the Object Parallax matches with the

expected visual perception norms.

7) Miscellaneous techniques

By modifying the perspective of an object within an image and by enhancing many of the minor depth cues the stereoscopic effect can be emphasised. The techniques below all operate using the 'Cut and Paste' technique. That is, a foreground object is 'Cut', enhanced and then 'Pasted' back on to the background.

- a) Shadows - Shading gives an object perspective.
- b) Foreground/Background - - By defocussing the background, through blurring or fogging, a foreground object may be emphasised, while defocussing the foreground object the background depth may be emphasised
- c) Edge Enhancement - Edges help to differentiate an object from its background.
- d) Texture Mapping - Helps to differentiate the object from the background.

LAYER 4 - 3D MEDIA (TRANSMISSION & STORAGE)

As for layer 1, layers 4 and 5 are not essential to the present invention. Layer 4 provides for the transmission and/or storage of the stereoscopic images. The transmission means can be adapted for a particular application. For example the following can be employed:

- 1) Local Transmission - can be via coax cable
- 2) Network TV Transmission - can be via
 - i) Cable
 - ii) Satellite
 - iii) Terrestrial
- 25 3) Digital Network - INTERNET, etc
- 4) Stereoscopic (3D) Image Storage
An image storage means may be used for storage of the image data for later transmission or display and may include :-
 - i) Analogue Storage - Video Tape, Film, etc
 - ii) Digital Storage - Laser Disk, Hard Disk, CD-ROM, Magneto Optical Disk, DAT, Digital Video Cassette (DVC), DVD.

LAYER 5 - 3D DISPLAY

As for the transmission means the display means can be dependent on the application requirements and can include:

1) Set-top Box

5 A set-top box by definition is a small box of electronics that receives, decodes, provides accessories interfaces and finally has outputs to suit the application. It may incorporate the following : -

- a) Video or RF receiver
- b) Stereoscopic (3D) decoder to provide separate left and right

10 10 image outputs to Head Mounted Devices or other stereoscopic displays where separate video channels are required.

- c) Resolution Enhancement - Line Doubling/Pixel Interpolation
- d) Shutter or Sequential Glasses Synchronisation
- e) Stereoscopic depth sensitivity control circuitry

15 f) Accessories interface - remote control with features such as a 2D/3D switch and Depth control.,

- g) Audio interface - audio output, headphone connection
- h) Access channel decoding - cable and pay TV applications
- i) Video or RF outputs

20 2) Stereoscopic Displays

Use special glasses or head gear to provide separate images to the left and right eyes including: -

- a) Polarising glasses - Linear and Circular polarisers
- b) Anaglyphic glasses - Coloured lenses - red/green,etc

25 c) LCD Shutter glasses

- d) Colour Sequential Glasses
- e) Head Mounted Devices (HMD) - Head gear fitted with two miniature video monitors (one for each eye), VR headsets

3) Autostereoscopic Displays

30 a) Video Projector/Retroreflective screen based display systems

- b) Volumetric display systems
- c) Lenticular lens based display systems

d) Holographic Optical Element (HOE) based display systems

PREFERRED EMBODIMENT

In summary, the present invention provides in a preferred embodiment a system that is capable of inputting monoscopic image sequences in a digital 5 format, or in an analogue format in which case an analogue to digital conversion process is involved. This image data is then subjected to a method of image analysis whereby the monoscopic images are compressed, if this is required for the particular application.

By comparing blocks of pixels in an image, with corresponding blocks in 10 an adjacent image, and by obtaining the minimum Mean Square Error for each block, motion within the image can be determined.

Following motion detection, regions of an image are identified for similar characteristics, such as, image brightness, colour, motion, pattern and edge continuity. The data is then subjected to motion analysis in order to determine 15 the nature of the motion in the image. This motion analysis takes the form of determining the direction, speed, type, depth and position of any motion in the image. This motion is then categorised into a number of categories including whether the motion is a complete scene change, a simple pan, a complex pan, an object moving on a stationary background, a stationary object in front of a 20 moving background, or whether there is no motion at all. Further actions are then determined based on these categories to convert the monoscopic images into stereoscopic image pairs suitable for viewing on an appropriate stereoscopic display device.

In the preferred embodiment, once the monoscopic images are analysed, 25 if a scene change or a complex pan is detected then no further analysis of that particular scene is required, rather the Field Delay and Field Delay history are both reset to zero. An object detection process is then applied to the new scene in order to try and identify objects within that scene. Once these objects are identified, then object processing takes place. If no objects are identified, then 30 the image is passed on for further processing using forced parallax and 3D optimisation.

If the motion categorised during the image analysis is not a scene

change, then further analysis of that scene is required. If further analysis of that scene results in the motion being categorised as a simple pan, then it is necessary to apply a Field Delay in accordance with the principles of motion parallax. It is then passed on for further processing. If the motion is not 5 categorised as a simple pan, but rather as an object in motion on a stationary background, then again we have to apply a Field Delay in accordance with the principles of motion parallax. In this regard, once the motion parallax has been applied, it is necessary to consider whether the objects all have a uniform direction. If the objects do move in a uniform direction, then it is passed on for 10 further processing at a later stage. If the objects do not have a uniform direction, then it is necessary to perform further object processing on selected objects within that scene to correct for the Reverse 3D effect. This can be achieved through using mesh distortion and morphing techniques.

If the motion is categorised as being a stationary object on a moving 15 background, it is then necessary to consider whether the background has a large variation in depth. If it does not, then we apply a Field Delay with the object having priority using the principles of motion parallax. However, if the background does have large variation in depth, then we apply a Field Delay with the background having priority as opposed to the object, again using the 20 principles of motion parallax. In this case, it is then also necessary to perform further object processing on the foreground object to correct for the Reverse 3D effect prior to being passed on for further processing.

If no motion is detected, then we next consider whether an object in the scene was known from any previous motion. If this is so, then we perform object 25 processing on that selected object. If not, then we apply an object detection process to that particular scene in order to attempt to identify any objects in it. If an object is identified, then we perform object processing on that particular object, if not, Forced Parallax and 3D Optimisation is performed.

Where object processing is required, objects are identified, tagged and 30 tracked, and then processed by using techniques of mesh distortion and morphing, object barreling, edge enhancement, brightness modification and object rotation.

In all cases, once the motion has been categorised and the primary techniques to convert to stereoscopic images have been applied, then a further amount of parallax or lateral shifting called forced parallax is applied to the image. It is noted that in the preferred embodiment, forced parallax is applied to
5 every image, not just for depth smoothing purposes but to provide an underlying stereoscopic effect that all images are seen as having depth behind or in front of the stereoscopic display device's reference plane, generally the front of the monitor screen. The advantages of applying forced parallax are that the system is better able to cope with changes in the category of the motion detected
10 without causing sudden changes in the viewers depth perception.

Once the forced parallax has been applied to the image, the image is then passed for 3D Optimisation. Again, this is not necessary in order to see a stereoscopic image, however the optimisation does enhance the image's depth perception by the viewer. The 3D Optimisation can take in a number of forms
15 including the addition of reference points or borders, parallax modulation, parallax smoothing and parallax adjustment for altering the depth sensitivity of any particular viewer. The image can also be optimised by modifying luminance or chrominance values pseudo randomly so that background motion behind foreground objects can be observed so that the depth perception is enhanced.
20 It is also possible to analyse for Reverse 3D so that a viewers eyestrain is minimised. Further techniques such as shadowing, foreground and background fogging or blurring and edge enhancement of the image can also be carried out in this stage.

Once the image has been optimised it is then transmitted to the
25 appropriate display device. This transmission can take a number of forms including cable, co-axial, satellite or any other form of transmitting the signal from one point to another. It is also possible that the image could be stored prior to being sent to a display device. The display device can take on a number of forms, and only need be appropriate for the application in hand, for example, it
30 is possible to use existing video monitors with a set top device in order to separate the left and right images, increase the scan rate and to synchronise viewing glasses. Alternatively, dedicated stereoscopic displays can be used

which incorporate the use of glasses or head gear to provide the stereoscopic images or alternatively, an auto-stereoscopic display device can be used. It is envisaged that the present invention will have application in theatres, cinemas, video arcades, cable or network TV, in the education area, particularly in the 5 multimedia industry and in many other areas such as theme parks and other entertainment applications.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An image conversion system for converting monoscopic images for viewing in three dimensions including:
 - an input means adapted to receive monoscopic images;
 - a preliminary analysis means to determine if there is any continuity between a first image and a second image of the monoscopic image sequence;
 - a secondary analysis means for receiving monoscopic images which have a continuity, and analysing the images to determine at least one of the speed and direction of motion, or the depth, size and position of objects;
 - a first processing means for processing the monoscopic images based on data received from the preliminary analysis means and/or the secondary analysis means.
2. An image conversion system as claimed in claim 1 further including a transmission means capable of transferring the processed images to a stereoscopic display system or a storage system.
3. An image conversion system as claimed in claim 1 or claim 2 wherein said first processing means processes the images by using at least one of:
 - motion parallax, forced parallax, parallax zones, image rotation or object processing.
4. An image conversion system as claimed in any preceding claim wherein a second processing means is provided to further process the images received from said first processing means.
5. An image conversion system as claimed in claim 4 wherein said second processing means uses forced parallax to further process the image.

6. An image conversion system as claimed in any one of claims 1 to 5 wherein a third processing means is provided for optionally enhancing the images prior to transmitting the converted images to the stereoscopic display device.
7. An image conversion system as claimed in claim 6 wherein said third processing means enhances the images by using at least one of:
reference points, parallax adjustment, parallax smoothing, parallax modulation, movement synthesis, reverse 3D correction or cut and paste techniques.
8. A method of manipulating monoscopic images into stereoscopic image pairs, including the steps of:
analysing the monoscopic images to determine the nature of motion within an image;
comparing any detected motion with a predefined range of motion categories;
processing the monoscopic images using at least one processing method depending on the motion category to form stereoscopic image pairs.
9. A method as claimed in claim 8 wherein said processing methods include motion parallax, forced parallax, parallax zones, image rotation and/or object processing.
10. A method as claimed in claim 8 or claim 9 wherein said motion categories include scene change, simple pan, complex pan, moving object, moving background and no motion.
11. A method as claimed in any one of claims 8 to 10 wherein the monoscopic image is compressed prior to any analysis.

12. A method as claimed in any one of claims 8 to 11 wherein the monoscopic image is scaled prior to any analysis.

13. A method as claimed in claim 12 wherein the scaling factor is dependent on the digital video resolution of each image .

14. A method for converting monoscopic images for viewing in three dimensions including:

a first layer adapted to receive a monoscopic image;

a second layer adapted to receive the monoscopic image and analyse the monoscopic image to create image data;

a third layer adapted to create stereoscopic image pairs from the monoscopic image using at least one predetermined technique selected as a function of image data;

a fourth layer adapted to transfer the stereoscopic image pairs to a stereoscopic display means;

a fifth layer consisting of a stereoscopic display means.

15. A method as claimed in claim 14 wherein said first layer is further adapted to convert any analogue images into a digital image.

16. A method as claimed in claim 14 or 15 wherein said second layer is adapted to detect any objects in a scene and make a determination as to the speed and direction of motion of any such objects.

17. A method as claimed in any one of claims 14 to 16 wherein the image is compressed prior to any analysis.

18. A method as claimed in any one of claims 14 to 17 wherein the third layer further includes an optimisation stage to further enhance the stereoscopic image pairs prior to transmitting the stereoscopic image pairs to the stereoscopic display means.

19. A method as claimed in any one of claims 14 to 18 wherein the fourth layer also includes a storage means for storing the stereoscopic image pairs for display on the stereoscopic display means at a later time.
20. A stereoscopic system including:
 - an input means capable of receiving monoscopic images;
 - an analysis means for analysing the images to determine an image category;
 - a conversion means capable of converting the monoscopic images into stereoscopic image pairs as a function of the selected image category for stereoscopic viewing.
21. A system as claimed in claim 20 wherein said input means includes a means to capture and digitise said monoscopic images.
22. A system as claimed in claim 20 or 21 wherein the image analysis means is capable of determining the speed, and direction of motion, of objects and the background within an image.
23. A system as claimed in any of claims 20 to 22 wherein the image analysis means is capable of determining the depth, size and position, of objects and the background within an image.
24. A system as claimed in any one of claims 20 to 23 further including a means for optimising the stereoscopic image to further improve the stereoscopic effect.
25. A system as claimed in any one of claims 20 to 24 further including a method of improving stereoscopic image pairs by adding a viewer reference point to the image.

26. A method of analysing monoscopic images for conversion to stereoscopic image pairs including the steps of:
 - scaling each image into a plurality of regions;
 - comparing each region of a first image with corresponding and adjacent regions of a second image to determine the nature of movement between said first image and said second image.
27. A method as claimed in claim 26 wherein a motion vector is defined for each image based on a comparison of the nature of motion detected with predefined motion categories ranging from no motion to a complete scene change.
28. A stereoscopic display system including the provision of a viewer reference point.
29. A system for converting monoscopic images for viewing in three dimensions including:
 - an input means adapted to receive monoscopic images;
 - a first analysis means for determining characteristics of the images;
 - a first processing means for processing the images based on the characteristics determined in said first analysis means;
 - an output means capable of transferring processed images to suitable storage and/or stereoscopic display systems.
30. A system as claimed in claim 29 wherein said input means is further adapted to digitise said monoscopic images.
31. A system as claimed in claim 29 or claim 30 further including a compression means adapted to compress said monoscopic images prior to analysis by said first analysis means.

32. A system as claimed in any one of claims 29 to 31 further including a scaling means adapted to scale said monoscopic image prior to analysis by said first analysis means.
33. A system as claimed in claim 32 wherein the scaling factor by which said monoscopic image is scaled is dependent on the digital video resolution of each image.
34. A system as claimed in any one of claims 29 to 33 further including:
a preliminary analysis means to determine if there is any continuity between successive first and second images.
35. A system as claimed in any one of claims 29 to 34 wherein said first analysis means is capable of determining objects within said images.
36. A system as claimed in any one of claims 29 to 35 wherein said first analysis means is capable of determining the motion of the images and/or the motion of objects within the images.
37. A system as claimed in claim 35 or claim 36 wherein said first analysis means is capable of categorising the motion into one of a predetermined range of motion categories.
38. A system as claimed in any one of claims 29 to 37 further including:
a second processing means for further processing the images.
39. A system as claimed in claim 38 wherein said second processing means uses forced parallax to further process the image.
40. A system as claimed in any one of claims 29 to 39 further including:
an optimisation means for enhancing the processed images prior to transferring the images to the stereoscopic display and/or storage system.

41. A system as claimed in any one of claims 29 to 40 further including:
a means to control the level of depth added to said monoscopic images.
42. A system as claimed in any one of claims 29 to 41 further including a
means to add a reference point to the processed image.
43. A method for converting monoscopic images for viewing in three
dimensions including the steps of:
receiving said monoscopic images;
analysing said monoscopic images to determine characteristics of the
images;
processing said monoscopic images based on the determined image
characteristics;
outputting the processed images to suitable storage and/or stereoscopic
display systems.
44. A method as claimed in claim 43 wherein said monoscopic image is
digitised before any analysis or processing is performed.
45. A method as claimed in claim 43 or claim 44 wherein said monoscopic
image is compressed prior to any analysis.
46. A method as claimed in any one of claims 43 to 45 wherein said
monoscopic image is scaled prior to any analysis.
47. A method as claimed in claim 46 wherein the scaling factor by which said
monoscopic image is scaled is dependent on the digital video resolution of each
image.
48. A method as claimed in any one of claims 43 to 47 wherein
successive first and second images are analysed for continuity before
determining the image characteristics.

49. A method as claimed in claim 48 wherein continuity is determined by comparing median luminance values between successive first and second images.

50. A method as claimed in claim 49 wherein no continuity is assumed when the difference in median luminance values exceeds 30.

51. A method as claimed in any one of claims 48 to 50 wherein the top few lines of successive images are compared to assist in the determination of continuity.

52. A method as claimed in any one of claims 43 to 51 wherein processing of images where no continuity is determined includes introducing a field delay to one eye such that the image which lacks continuity is seen by one eye prior to being viewed by the other eye of a viewer.

53. A method as claimed in any one of claims 43 to 52 wherein during analysis of said monoscopic images, objects within said images are defined to assist during said processing.

54. A method as claimed in any one of claims 43 to 53 wherein during analysis of said monoscopic image the motion of the images and/or objects within the images is determined to assist said processing.

55. A method as claimed in claim 54 wherein the motion of said image and/or objects is categorised into one of a predetermined range of motion categories.

56. A method as claimed in any one of claims 43 to 55 wherein analysing of said monoscopic images to determine the motion includes the steps of:
dividing each image into a plurality of blocks, wherein corresponding blocks on an adjacent image are offset horizontally and/or vertically;

comparing each said block with said corresponding blocks to find the minimum mean square error and thereby the motion of the block.

57. A method as claimed in claim 56 wherein any said block with no details is not compared with said corresponding blocks.

58. A method as claimed in claim 56 when appended to claim 48 wherein no continuity is assumed when the comparison of the majority of blocks with said corresponding blocks has resulted in large error values.

59. A method as claimed in any one of claims 43 to 58 wherein processing of the image includes the use of motion parallax by introducing a field delay such that one eye of a viewer views the image before the other eye of the viewer.

60. A method as claimed in claim 59 wherein the amount of motion is inversely proportional to the field delay.

61. A method as claimed in claim 59 or claim 60 wherein the field delays are stored, and the field delay for each new image is averaged against previous field delays.

62. A method as claimed in claim 61 wherein stored field delays are deleted when a non continuity is detected.

63. A method as claimed in any one of claims 43 to 62 wherein processing of the image includes the use of forced parallax by introducing a lateral shift through displacement of the left and right eye images.

64. A method as claimed in any one of claims 43 to 63 wherein processing of the image includes the use of parallax zones by introducing a greater lateral shift to one portion of the image.

65. A method as claimed in any one of claims 43 to 64 wherein processing of the image includes a combination of forced parallax and motion parallax on various parts of the image.

66. A method as claimed in any one of claims 43 to 65 wherein processing of the image includes rotation of the left and right eye images about the y axis an equal amount in an opposite direction.

67. A method as claimed in any one of claims 43 to 66 wherein processing of the image includes the use of at least one of the following object processing techniques:

- mesh distortion and morphing
- object barreling
- object edge enhancement
- object brightness enhancement
- object rotation.

68. A method as claimed in any one of claims 43 to 67 wherein the processed image is further processed by applying a final forced parallax to the processed image.

69. A method as claimed in claim 68 wherein the degree of forced parallax is determined by the amount of parallax added during processing of the image, such that the total of the parallax added during processing and the forced parallax, is substantially equal to the total parallax of adjacent images.

70. A method as claimed in claim 68 or claim 69 wherein the degree of final forced parallax is modulated between predetermined minimum and maximum settings over a predetermined time frame.

71. A method as claimed in any one of claims 43 to 70 wherein the processed image is optimised to further enhance the processed images prior to transferring the images to the stereoscopic display and/or storage system.

72. A method as claimed in any one of claims 43 to 71 wherein a reference point is added to the processed image.

73. A method as claimed in claim 72 wherein said reference point is at least one of:

- a border around the perimeter of the image,
- a plurality of concentric borders,
- a partial border,
- a logo,
- a picture.

74. A method as claimed in any one of claims 43 to 73 wherein the amount of depth added to the monoscopic images during processing of the images can be adjusted in response to a viewers preference.

75. A method as claimed in any one of claims 43 to 74 wherein the background of the image is randomly moved in small increments which are not consciously noticed by the viewer.

76. A method as claimed in any one of claims 43 to 75 wherein the image is tested for reverse 3D and objects manipulated individually to compensate for any reverse 3D.

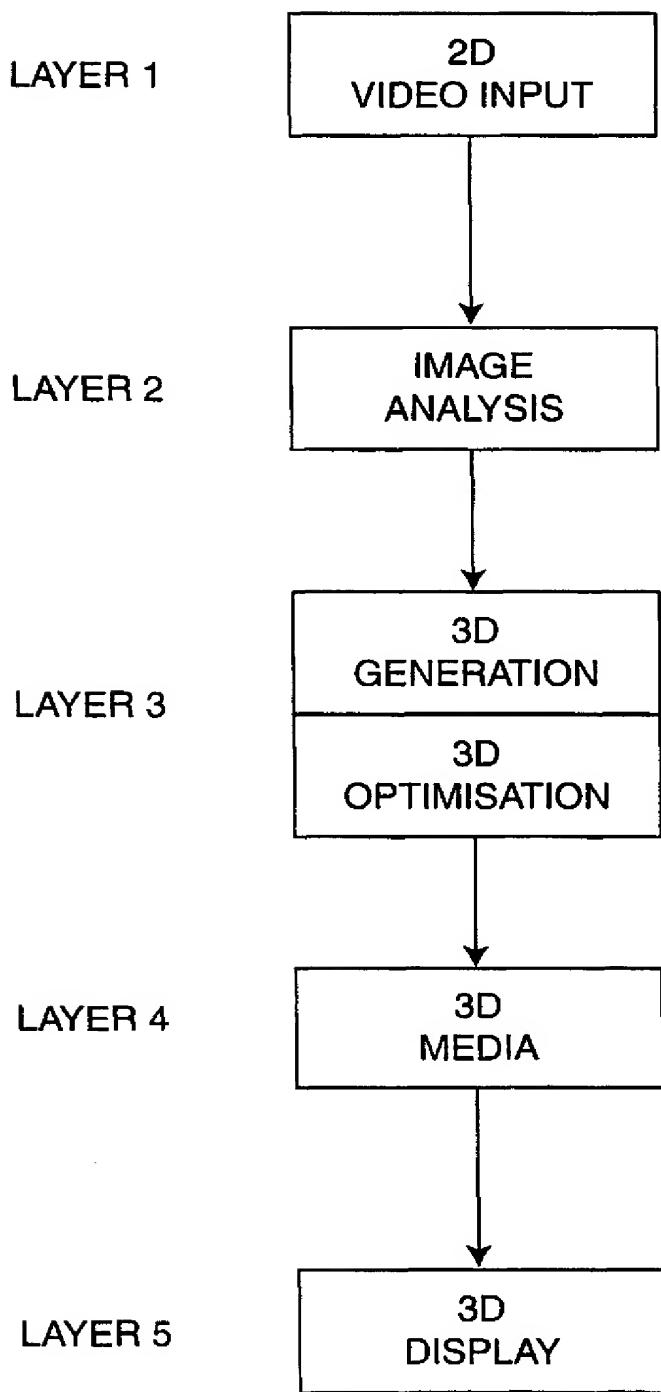
77. A method as claimed in any one of claims 43 to 76 wherein cut and paste techniques are employed to further emphasise the stereoscopic effect.

78. A method substantially as hereinbefore disclosed with reference to the accompanying drawings.

79. A system substantially as hereinbefore disclosed with reference to the accompanying drawings.

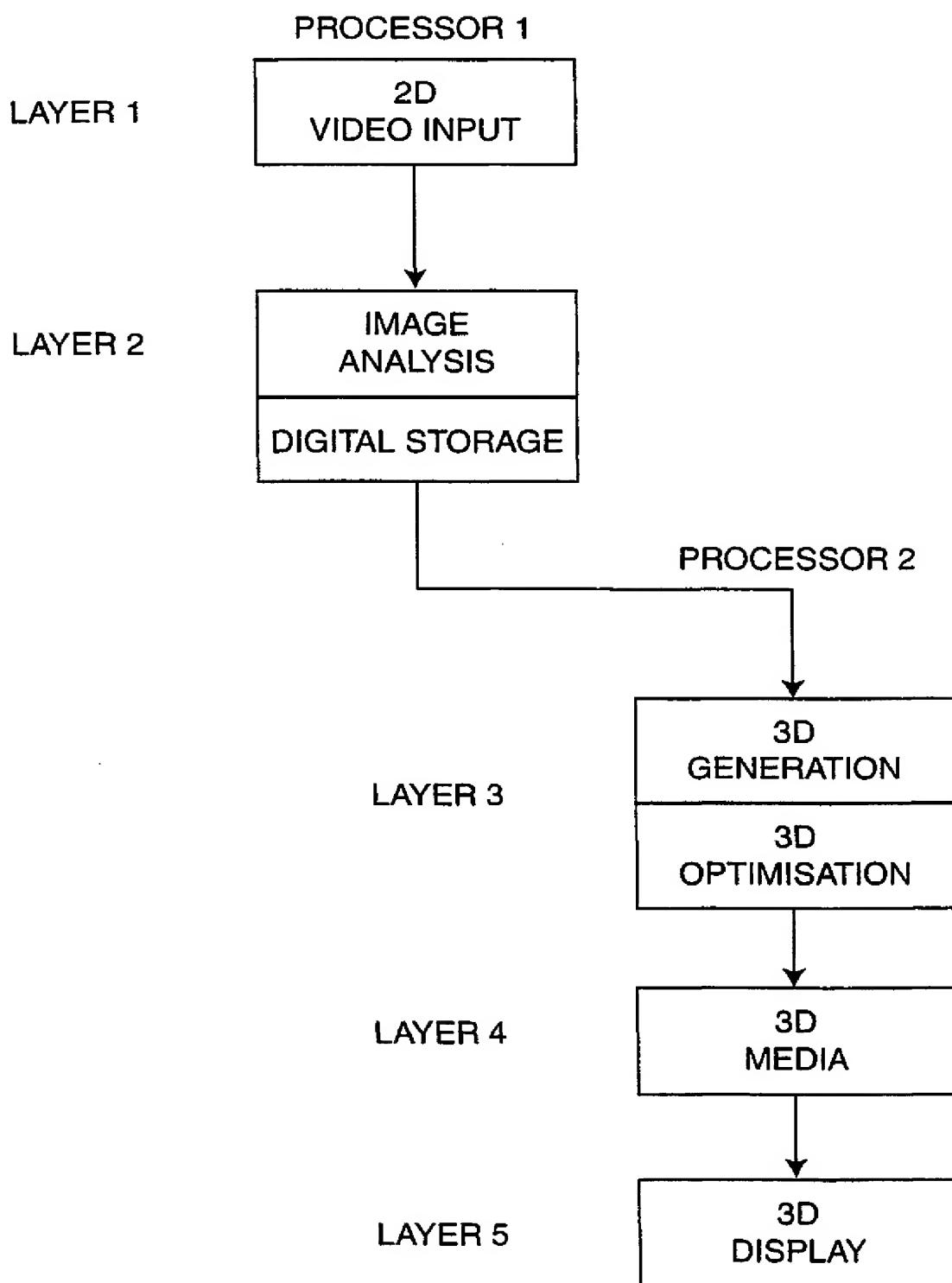
1/8

Figure 1.



2/8

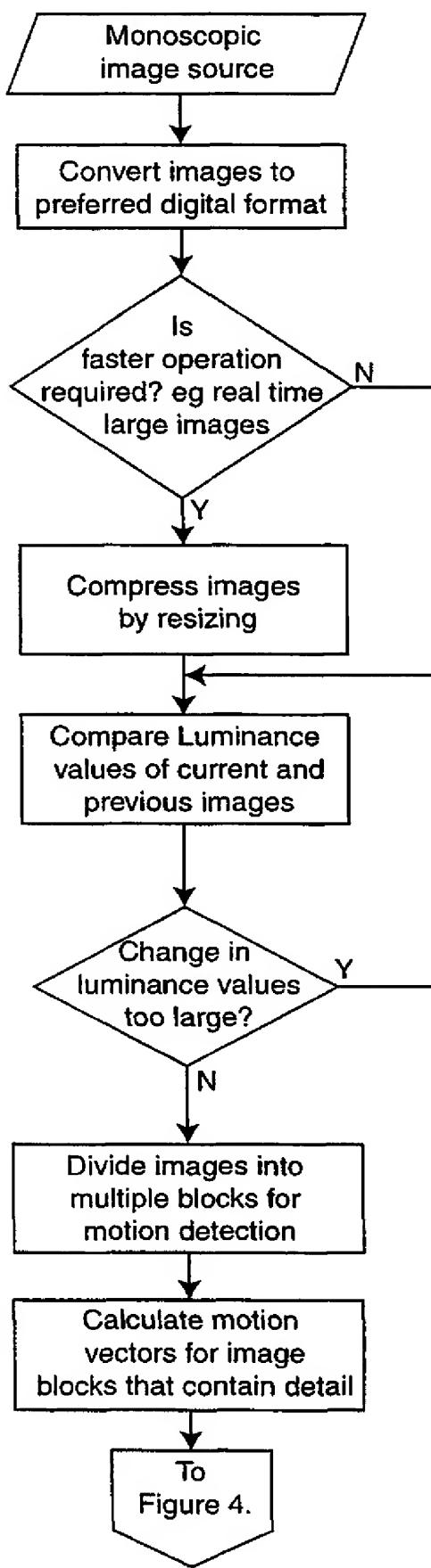
Figure 2.

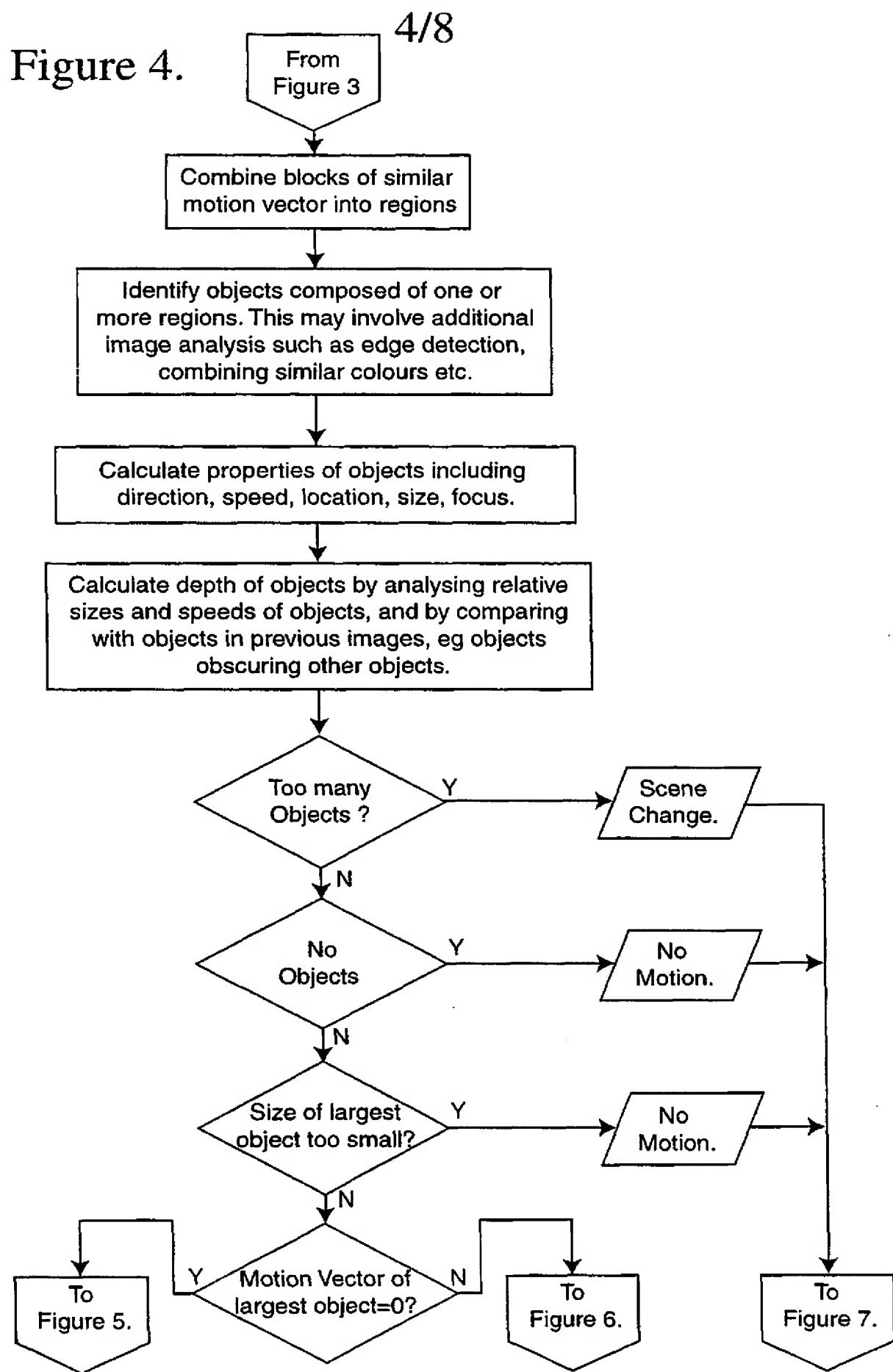


3/8

Figure 3.

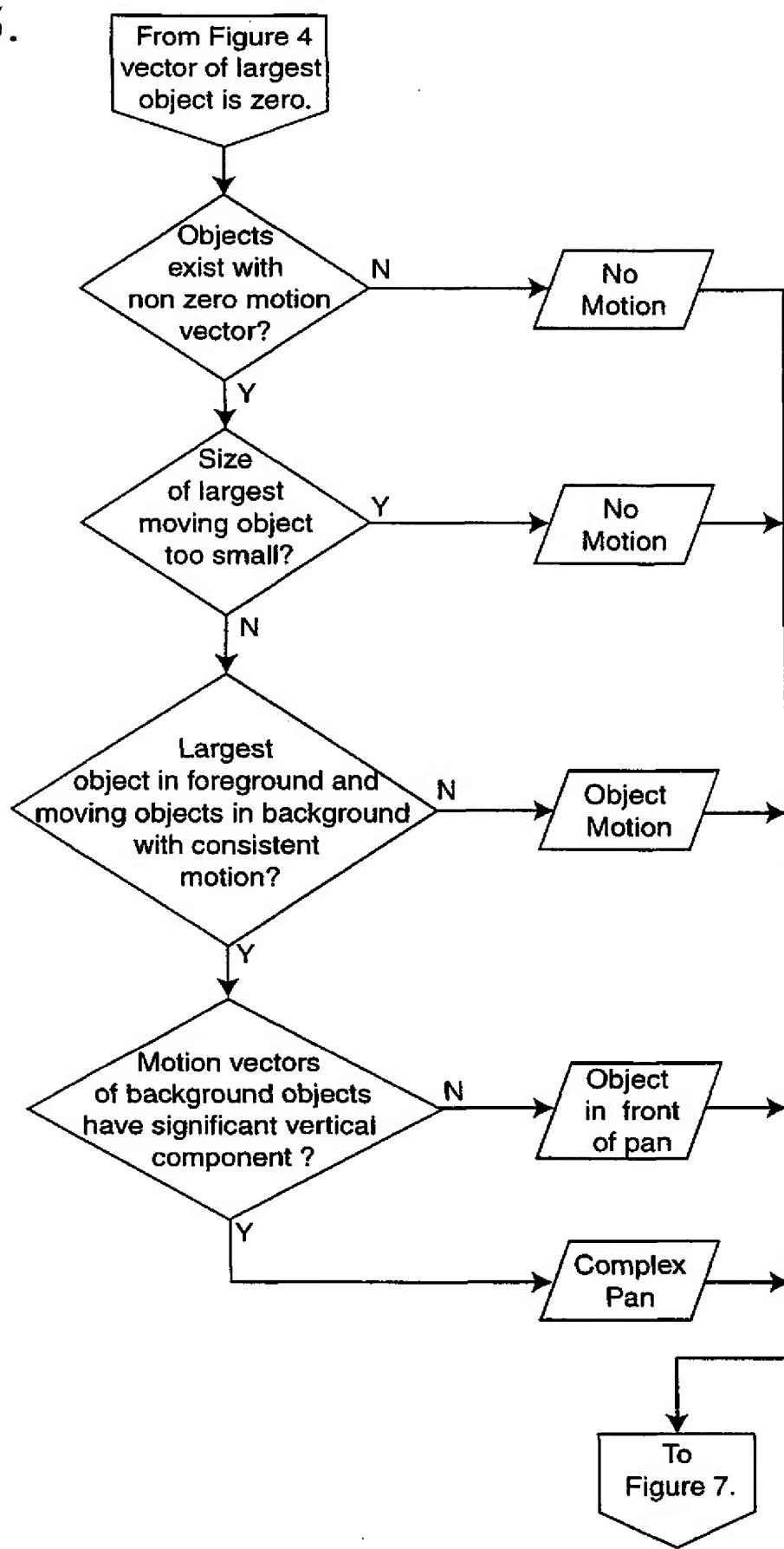
Layer 1





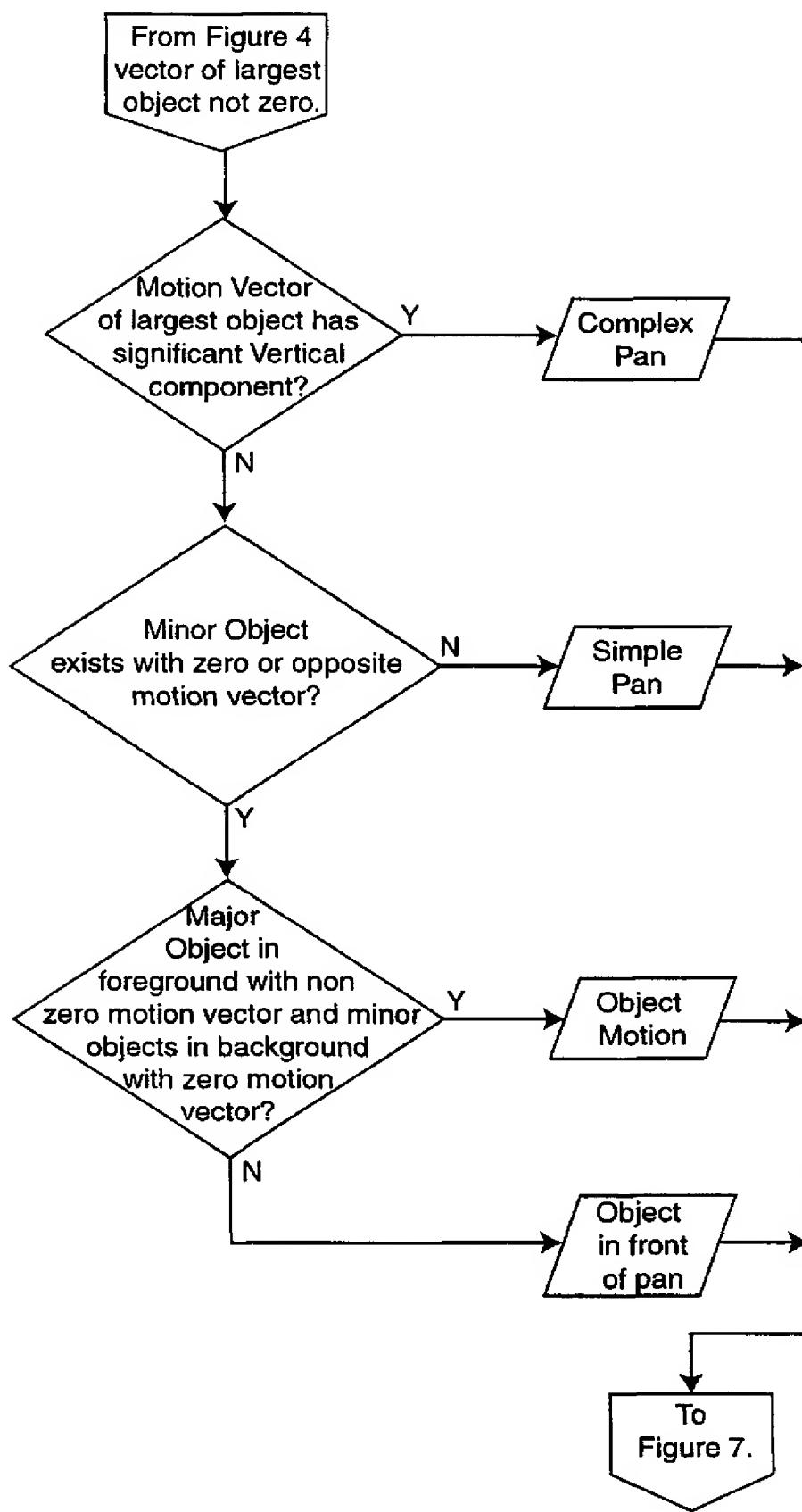
5/8

Figure 5.



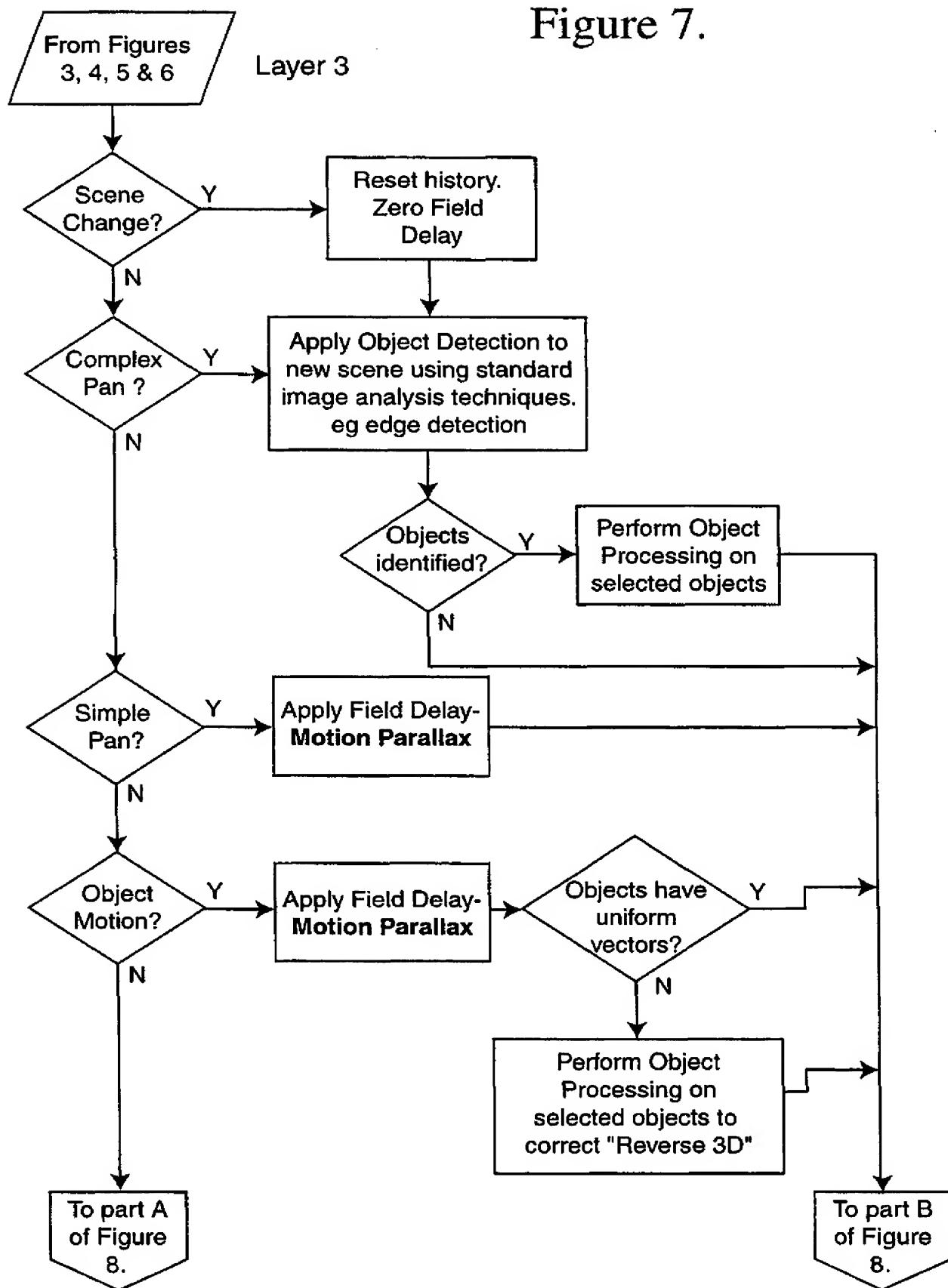
6/8

Figure 6.



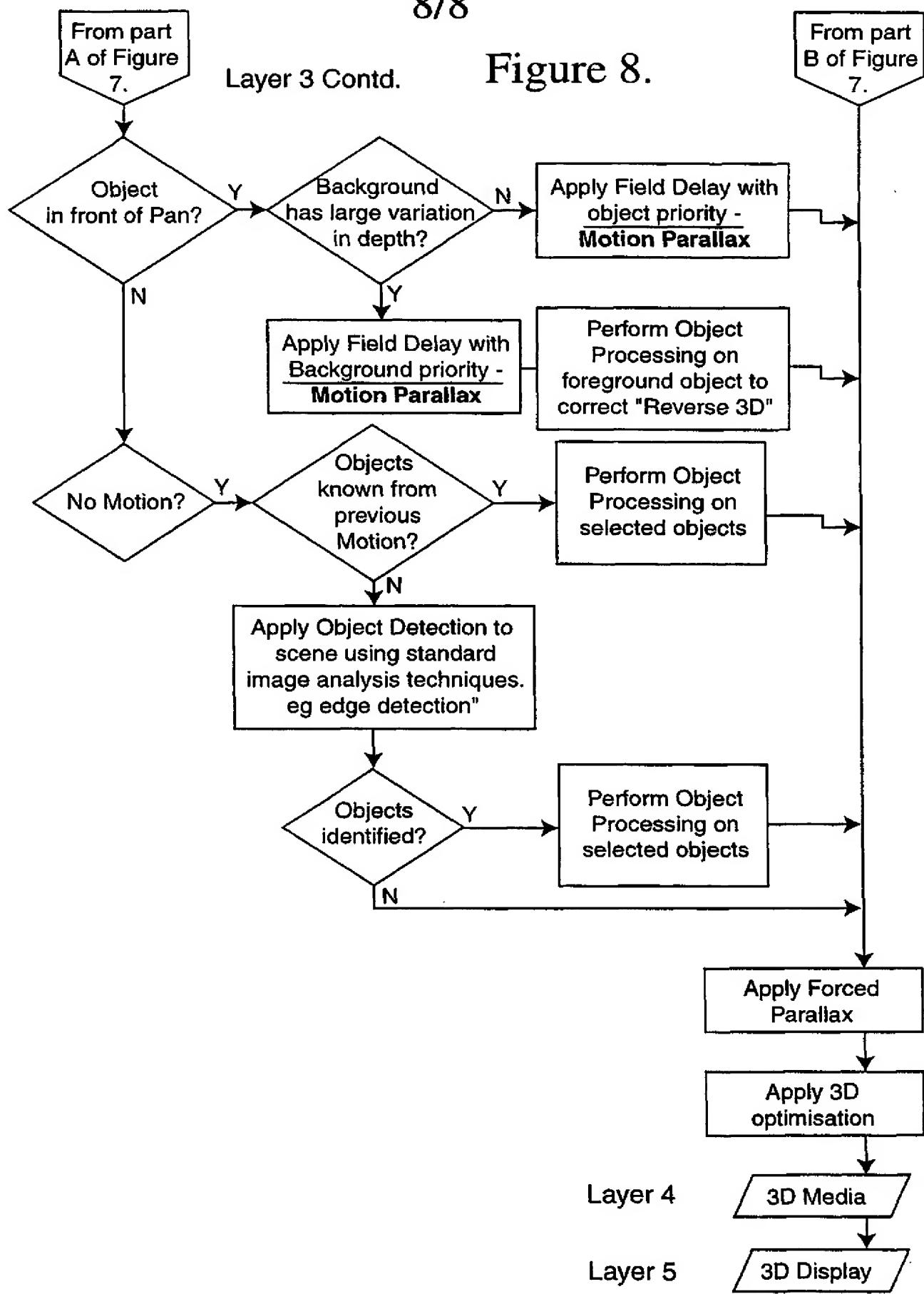
7/8

Figure 7.



8/8

Figure 8.



INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 98/00716

A. CLASSIFICATION OF SUBJECT MATTERInt Cl⁶: G06T 15/00, H04N 13/00, 15/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06T 15/00, H04N 13/-, 15/00, 9/54, 9/60, G03B 35/-, G01C 11/04, 11/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

AU: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 WPAT: IPC as above with keywords (STEREOSCOP: or THREE DIMENSION: or 3D) and IMAGE: and (CONVER: or
 JAPIO: TRANSFORM: or MANIPULAT: or CHANG: or TRANSPOS: or ALTER:) and (MONOSCOP: or
 MONOCUL: or TWO DIMENSION: or 2D) and (ANALYS: or PROCESS: or COMPAR: or SCAL: or
 RESOLV: or OPERAT: or DIVID:)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AU 14635/97 A (SYTHONICS INC) 14 July 1997 page 26 line 11-page 29 line 27, page 38 line 30-page 44 line 31, page 51 line 1- page 52 line 24, page 53 line 26-page 54 line 19, page 68 line 20-36, page 94 line 9-page 95 line 18, page 108 line 19-page 109 line 2, page 114 line 1-page 115 line 4, Figs. 1-6, 15, 17, 19, 62, 75, 80	1-7,14,15,17-21,23,24, 26,29-35,38-41, 43-48,53,63,64, 67-71,74,77
P,X	US 5777666 A (TANASE et al) 7 July 1998 col 6 line 22-col 8 line 4, Fig 1.	14-16,29,30,35- 38,43,44,53,54,59-61

Further documents are listed in the continuation of Box C

See patent family annex

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 8 October 1998	Date of mailing of the international search report 16 OCT. 1998
---	--

Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200 WODEN ACT 2606 AUSTRALIA Facsimile No.: (02) 6285 3929	Authorized officer M.E. DIXON Telephone No.: (02) 6283 2194
--	---

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 98/00716

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	US 5760933 A (ARITAKE et al.) 2 June 1998 col 8 line 13-21, col 8 line 43-col 9 line 17, Figs. 9, 12, 13	26,29,32,35,36,38-43, 46,53,54,64,71-73
P,X	US 5717415 A (IUE et al.) 10 February 1998 col 4 line 64-col 7 line 67, Figs 3, 4	14-16,29,30, 35-38,43,44,53,54, 59-61
P,X	US 5682437 A (OKINO et al.) 28 October 1997 whole specification	1-10,12-16,18-22,24, 26,27,29,30,32-41, 43,44,46-48,52-55, 59-65,68,71,76
X	US 5510832 A (GARCIA) 23 April 1996 whole specification	14,15,17-19
X	US 4925294 A (GESHWIND et al.) 15 May 1990 see in particular col 4 line 52-58, col 5 line 28-47	14-16,18-21, 23-25,29,30,35,36, 38,40-44,53,54,64, 67,71-75,77
X	WO 97/24000 A (XENOTECH RESEARCH PTY LTD) 3 July 1997 pages 1-14, Figs 1-8	14,15,1921,23,29,30, 35,38,41,43,44,53,64, 66,67,77
X	WO 96/41311 A (GESHWIND) 19 December 1996 whole specification	14,15,19-21,23,24, 29,30,35,38-41,43, 44,53,63,66-68,71, 74,75
X	EP 0704823 A (SANYO ELECTRIC CO LTD) 3 April 1996 page 6, lines 3-9, page 7 line 14-page 8 line 32, Figs 1-5	1-4,6-10,12, 14-16,26,27,29,30, 34-36,38,43,44,46,48, 53,54,59-61

INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU 98/00716

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: 78,79
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claims 78, 79 rely on reference to description and drawings (Rule 6.2(a))

3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. Claims 1-27, 29-77 directed to systems and methods involving analysing monoscopic images for conversion to stereoscopic. It is considered that the analysis of monoscopic images for such conversion comprises a first "special technical feature".
2. Claim 28 directed to a stereoscopic display system including the provision of a viewer reference point. It is considered that the inclusion of a viewer reference point in a stereoscopic display system comprises a second "special technical feature".

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-27, 29-77

Remark on Protest

The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/AU 98/00716

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report			Patent Family Member				
AU	14635/97	WO	97/22951	US	5748199		
US	5777666	JP	9009295	JP	9037300		
US	5760933	CA	2100109	EP	580354	JP	6035391
US	5717415	CN	1126344	EP	665697	JP	7287773
		JP	9107562	JP	7226958		
US	5682437	EP	703716	JP	8149512	JP	8149513
		US	5699443	EP	704823	JP	8149517
US	5510832	AU	12904/95	CA	2177051	EP	746949
		WO	95/15661				
US	4925294	AU	10857/88	EP	442865	WO	88/04804
WO	97/24000	AU	10884/97	EP	868818		
WO	96/41311	AU	61082/96				
EP	704823	JP	8149517	US	5699443	US	5682437
		EP	703716	JP	8149513	JP	8149512

END OF ANNEX